A SAVAGE-LIKE AXIOMATIZATION FOR NONSTANDARD EXPECTED UTILITY

GRANT MOLNAR

Abstract. Since Leonard Savage's epoch-making "Foundations of Statistics" [13], Subjective Expected Utility Theory has been the presumptive model for decision-making. Savage provided an act-based axiomatization of standard expected utility theory. In this article, we provide a Savage-like axiomatization of nonstandard expected utility theory. It corresponds to a weakening of Savage's 6^{th} axiom.

1. Introduction

In the last twenty years, there has been an explosion of research in decision theory, and various decision-making algorithms have been proposed as descriptive or normative alternatives to expected utility theory, such as info-gap theory ([2]), Choquet expected utility ([1], [3]), qualitative binary possibilistic utility ([6], [14]), and so forth. One approach that has garnered some attention is a refinement of expected utility theory by allowing utility functions and probability measures to contain infinitesimal or infinite elements. This refinement permits coherent modeling of Pascal's Wager ([10]), and allows uncountably many pairwise disjoint events to have nonzero probability, circumnavigating some of the troubles of standard expected utility theory ([9]). It has also been defended on game-theoretic grounds ([8]).

One of the most popular methods for modeling infinitesimals is Abraham Robinson's nonstandard analysis, which enriches the field of real numbers with infinitely small and infinitely large numbers in a way that sensibly preserves the field's underlying structure (for more details, see [7], [11]). This extension of the reals is termed the hyperreal numbers and generally denoted by $^*\mathbb{R}$.

After a brief overview of some pertinent results, we prove the main result of the paper

2. Nonstandard Analysis

We provide a brief survey of some elementary constructions and theorems from nonstandard analysis. No proofs are provided; our exposition is taken directly from chapters 2 through 4 of [7], so curious readers are encouraged to study further there.

Definition 2.1. *A nonprincipal ultrafilter on $\mathbb{N}$ is a set $\mathbb{U} \subseteq 2^{\mathbb{N}}$ such that*

1. $\forall A, B \in \mathbb{U}, A \cap B \in \mathbb{U}$ $\quad (filter)$
2. *If* $A \in \mathbb{U}, A \subseteq B \subseteq \mathbb{N},$ *then* $B \in \mathbb{U}$ $\quad (filter)$
3. $\emptyset \notin \mathbb{U}$ $\quad (ultrafilter)$
4. $\forall A \subseteq \mathbb{N}, A \in \mathbb{U}$ *or* $\mathbb{N} - A \in \mathbb{U}$ $\quad (ultrafilter)$
5. $\forall a \in \mathbb{N}, \exists A \in \mathbb{U} : a \notin A$ $\quad (nonprincipal)$

Brigham Young University, Provo.

Observation 2.1.1. *Every nonprinicipal filter is cofinite: if $A \subset \mathbb{N}$ has finite cardinality, then $\mathbb{N} - A \in \mathbb{U}$.*

Definition 2.2. *Fix $\mathbb{U}$ a nonprincipal ultrafilter. For S an arbitrary set, define ${}^*S$ to be $S^{\mathbb{N}}/\sim$, where $(s_1, s_2, ...) \sim (r_1, r_2, ...)$ if $\{i \in \mathbb{N} : r_i = s_i\} \in \mathbb{U}$.*

Observation 2.2.1. *$\sim$ is an equivalence relation on $S^{\mathbb{N}}$. Elements of S are embedded in ${}^*S$ diagonally by $s \mapsto (s, s, s, ...)$.*

Observation 2.2.2. *The nonstandard extension of a relation R on S is defined as follows: ${}^*rR{}^*s$ if, for $(r_1, ...), (s_1, ...)$ representatives of ${}^*r$ and ${}^*s$ respectively, $\{i \in \mathbb{N} : r_iRs_i\} \in \mathbb{U}$. This is well-defined.*

Observation 2.2.3. *The nonstandard extension of $f : R \to S$ is defined as follows: for $(r_1, ...)$ a representative of ${}^*r \in^* R$, let ${}^*f({}^*r)$ be the equivalence class containing $(f(r_1), f(r_2), ...)$. This is well-defined.*

Observation 2.2.4. *Similar extensions can be made of binary operations, cartesian products, and every other set-theoretic construction. All are well-defined.*

Example 1. *${}^*\mathbb{R} = \mathbb{R}^{\mathbb{N}}/\sim$ is a totally ordered field which contains infinitesimal and infinite elements (that is, $\exists \epsilon \in {}^*\mathbb{R} : \forall n \in \mathbb{N}, 0 < \epsilon < \frac{1}{n}$, and $\exists E \in {}^*\mathbb{R} : \forall n \in \mathbb{N}, n < E$). We say a function f mapping into ${}^*\mathbb{R}$ is hyperbounded if $\exists\ {}^*n \in {}^*\mathbb{N}$ such that $sup|f(x)| < {}^*n$. Note that a function may be hyperbounded without being bounded.*

Observation 2.2.5. *$\forall\ {}^*n \in {}^*\mathbb{N}$, $|\{\ {}^*k \leq {}^*n\ \}| \leq |\mathbb{N}|$, and this bound is sharp.*

Proof. Let $(n_1, n_2, ...)$ be a representative for ${}^*n$. Clearly, the set $A = \{(a_1, , ...) : a_i \leq n_i \forall i \in \mathbb{N}\}$of sequences bounded pointwise from above by $(n_1, n_2, ...)$ no more than countable, and when we identify elements of A modulo $\sim$ cannot increase. On the other hand, consider ${}^*n$ $\in {}^*\mathbb{N}$ containing $(1, 2, 3, 4, ...)$. As $\forall k \in \mathbb{N}$, $\{i \in \mathbb{N} : k \geq i\}$ is finite, ${}^*n \geq k\ \forall k \in \mathbb{N}$ and so $|\{\ {}^*k \leq {}^*n\ \}| = |\mathbb{N}|$. Then the bound is sharp, as desired. □

One of the most important theorems in nonstandard analysis is the transfer principle, which asserts that for any mathematical object S, its nonstandard extension ${}^*S$ meaningfully preserves its properties.

Theorem 2.1 (Transfer Principle)**.** *Let $\mathcal{R}$ be a relational structure, $\mathcal{L}_{\mathcal{R}}$ the mathematical language of $\mathcal{R}$ (comprising the relation and function symbols of $\mathcal{R}$ together with logical connectives, and existential qualifiers, and parentheses). A defined $\mathcal{L}_{\mathcal{R}}$ sentence ϕ is true if and only if ${}^*\phi$ is true.*

Roughly what this means is that any statement ϕ about a mathematical object S is true if and only if the nonstandard analogue of ϕ (constructed by extending the objects, relations, and functions in ϕ to ${}^*S$) is true.

As the Transfer Principle suggests, it is possible to develop nonstandard analysis on purely model-theoretic grounds, without any reference to ultrafilters, but we will not pursue that train of thought further here.

3. Standard Expected Utility Theory

Leonard Savage proved that any decision-maker who complied with seven plausible rationality axioms behaved as if they followed expected utility theory. We reproduce his axioms and theorem here, but make no effort to justify these axioms: interested readers are directed to [5] and [13].

First, we provide some definitions. The state space S is the space of possible states of the world. This space is assumed to be exhaustive and mutually exclusive, so that the world is identified with exactly one element of S. X is the space of possible outcomes, or results of the actor's choice. Let $D = S^X$ denote the space of conceivable decisions that the actor could undertake. In practice, an actor will normally only be able to choose from a small subset of D, but Savage considered it reasonable to assume a normatively rational actor would have the capacity to compare any two hypothetical decisions. This is not a restrictive assumption, as we may enrich the decision space with additional options without forcing any changes to the actor's original preferences. We make the further assumption that the actor's menu of choices and choice does not alter the state $s \in S$ of the world.

The actor is supposed to have preferences between possible decisions which are described by the binary relation "$\succ$." $f \succ g$ is read as "f is preferred to g." From this primitive binary relation, we may define $f \prec g$ to hold if $g \succ f$, $f \sim g$ if $f \not\succ g$ and $g \not\succ f$, and $f \succeq g$ if $f \succ g$ or $f \sim g$. For $x, y \in X$, we also say $x \succ y$ if the constant function returning x is preferred to the constant function returning y. At the outset, we make no assumptions about the structure of this preference relation.

We employ the following definitions:

Definition 3.1. *For $f, g \in D, A \subseteq S$, $fAg(s)$ is the function that returns $f(s)$ if $s \in A$, $g(s)$ else.*

Definition 3.2. *For $A, B \subseteq S$, $A \succ_L B$ (read "A is more likely than B") if whenever $x \succ y$, $xAy \succ xBy$.*

Definition 3.3. *For $A \subseteq S$, $(f \succ g)_A$ (read "f is preferred to g given A") if $\forall h \in D$, $fAh \succ gAh$.*

Definition 3.4. *$A \subseteq S$ is null if $\forall f, g \in D, (f \sim g)_A$*

Theorem 3.1 (Savage)**.** *Suppose the following conditions hold:*

- *S1.* *$\succ$ on D is a weak order.*
- *S2.* *For $f, g, f', g' \in D, A \subseteq S, (fAg \succ f'Ag) \Rightarrow (fAg' \succ f'Ag')$*
- *S3.* *For $x, y \in X, A$ not null, $((x \succ y)|_A \iff x \succ y)$*
- *S4.* *For $x \succ y$, $x' \succ y'$, $A, B \subseteq S$, $xAy \succ xBy \iff x'Ay' \succ x'By'$*
- *S5.* *$\exists x, y \in X$ such that $x \succ y$.*
- *S6.* *For all $f, g, h \in D$, if $f \succ g$, there exists $n \in \mathbb{N}$ and a partition $E_1, E_2, ...E_n$ of S such that for all $k \in \{1, ..., n\}, fE_kh \succ g$, and $f \succ gE_kh$.*
- *S7.* *If $(f(s) \succ g)|_A \forall s \in A$ then $(f \succeq g)|_A$. If $(f \succ g(s))|_A \forall s \in A$ then $(f \succeq g)|_A$.*

Then there is a unique finitely additive probability measure P on 2^S such that

$$\forall A, B \subseteq S,\ A \succ_L B \iff P(A) > P(B)$$

and P has the property that

$$\forall B \subseteq S, \rho \in [0,1],\ \exists C \subseteq B : P(C) = \rho P(B)$$

Furthermore, with P as given, there exists a function u from X to $\mathbb{R}$ for which

$$f \succ g \iff \int_S u(f(s))dP(s) > \int_S u(f(s))dP(s)$$

and this u is bounded and unique up to positive affine transformation.

Proof. Theorem 14.1 of [5] □

4. Main Result

We now prove our main result.

Theorem 4.1. *With all the notation of Savage's Theorem (3.1), let $\succ$ satisfy axioms $S1$, $S2$, $S3$, $S4$, $S5$, and $S7$, and the following modified version of $S6$:*

S6'. For all $f, g, h \in D$, if $f \succ g$, there exists an at most countable partition $E_1, E_2, ...$ of S such that for all $k \in \mathbb{N}$, $fE_kh \succ g$, and $f \succ gE_kh$.

Then there is a unique countably additive nonstandard probability measure P on 2^S such that

$$\forall A, B \subseteq S,\ A \succ_L B \iff P(A) > P(B)$$

and P has the property that

$$\forall B \subseteq S, \rho \in^* [0,1],\ \exists C \subseteq B : P(C) = \rho P(B)$$

Furthermore, with P as given, there exists a function u from X to ${}^*\mathbb{R}$ for which

$$f \succ g \iff \int_S u(f(s))dP(s) > \int_S u(f(s))dP(s)$$

and this u is hyperbounded and unique up to positive affine transformation.

Proof. Apply the transfer principle (2.1) to Savage's Theorem (3.1). Clearly, for axioms $S1 - S5$ and $S7$, this is only a relabeling of S as ${}^*S$, X as X. Axiom $S6$ becomes "For all $f, g, h \in D$, if $f \succ g$, there exists ${}^*n \in {}^*\mathbb{N}$ and a partition $E_1, E_2, ...E_{{}^*n}$ of S such that for all $k \in \{1, ..., {}^*n\}$, $fE_kh \succ g$, and $f \succ gE_kh$." But for any given ${}^*n \in {}^*\mathbb{N}$, there are at most countably many $k \leq {}^*n$. If ${}^*n$ happens to be finite, then clearly the partition $E_1, ..., E_{{}^*n}$ is at most countable, since it is finite. Else, since $|\{k \leq {}^*n\}| = |\mathbb{N}|$, there exists a bijection $\phi : \{k \leq {}^*n\} \to \mathbb{N}$, so we may relabel $\{E_k\}_{k \leq {}^*n}$ as $\{E_{\phi(k)}\}_{k \leq {}^*n} = \{E_k\}_{k \in \mathbb{N}}$. Note also that by observation 2.2.5, we know there are ${}^*n$ in ${}^*\mathbb{N}$ which correspond to a countable partition of S.

Observe also that as S, X have no assumed structure, neither do ${}^*S$ or ${}^*X$. Then we can simply relabel ${}^*S$ and ${}^*X$ as S and X without losing any information. Applying theorem 2.1 to the assertion $\forall n \in \mathbb{N}, A_1, ...A_n$ pairwise disjoint measurable sets, we have $P(\cup_{i=1}^{{}^*n} A_i) = \sum_{i=1}^{{}^*n} P(A_i)$ for some ${}^*n$ in ${}^*\mathbb{N}$. Then by observation 2.2.5, P is countably additive. The proposition follows. □

Corollary 4.2. *Suppose $\succ$ on D satisfies conditions $S1 - S5$, $S6'$, $S7$ above. Then we may take u to be bounded.*

Proof. As u is hyperbounded, $\exists n \in {}^*\mathbb{N}$ such that $\underset{x \in X}{sup}\ u(x) < n$ Then $\frac{u}{n}$ is a positive affine transformation of u which is bounded. □

5. Conclusion

Our axiomatization justifies a decision-theoretic model for individuals who generally agree with Savage's postulates but feel uncomfortable gambling extremely valuable goods (like their lives) against relatively unimportant goods (like a dollar) at even the slimmest of odds. It also makes clearer exactly what postulates a person would need to accept in order to accept nonstandard expected utility, highlighting the gap between Bayesians and decision theorists in the school of Ellsburg [4].

A variety of perspectives stress the advantages of working with nonstandard probability theory rather than its more granular standard counterpart. It is therefore natural to ask what advantages the nonstandard approach has for other decision theories. As many of

these theories have been axiomatized in a Savage-like manner (see for instance [12], [15]), such axiomatizations should be similarly easy to derive by the transfer principle (2.1). We leave this question for future researchers to investigate.

References

[1] Takao Asano and Hiroyuki Kojima. An axiomatization of Choquet expected utility with cominimum independence. *Theory and Decision*, 78(1):117–139, 2015.

[2] Yakov Ben-Haim. *Info-Gap Decision Theory: Decisions Under Severe Uncertainty*. Academic Press, 2006.

[3] Alain Chateauneuf, Jürgen Eichberger, and Simon Grant. A simple axiomatization and constructive representation proof for Choquet expected utility. *Econom. Theory*, 22(4):907–915, 2003.

[4] Daniel Ellsberg. *Risk, Ambiguity and Decision*. Routledge, 2016.

[5] Peter C Fishburn. Utility theory for decision making. Technical report, DTIC Document, 1970.

[6] Phan H. Giang and Prakash P. Shenoy. Two axiomatic approaches to decision making using possibility theory. *European J. Oper. Res.*, 162(2):450–467, 2005.

[7] Robert Goldblatt. *Lectures on the hyperreals*, volume 188 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1998. An introduction to nonstandard analysis.

[8] Peter J. Hammond. Elementary non-Archimedean representations of probability for decision theory and games. In *Patrick Suppes: scientific philosopher, Vol. 1*, volume 233 of *Synthese Lib.*, pages 25–61. Kluwer Acad. Publ., Dordrecht, 1994.

[9] Peter J. Hammond. Non-Archimedean subjective probabilities in decision theory and games. *Math. Social Sci.*, 38(2):139–156, 1999.

[10] Frederik Herzberg. Hyperreal expected utilities and Pascal's wager. *Logique et Anal. (N.S.)*, 54(213):69–108, 2011.

[11] Abraham Robinson. *Non-standard analysis*. Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 1996. Reprint of the second (1974) edition, With a foreword by Wilhelmus A. J. Luxemburg.

[12] Rakesh Sarin and Peter Wakker. A simple axiomatization of nonadditive expected utility. *Econometrica*, 60(6):1255–1272, 1992.

[13] Leonard J. Savage. *The foundations of statistics*. Dover Publications, Inc., New York, revised edition, 1972.

[14] Paul Weng. An axiomatic approach to qualitative decision theory with binary possibilistic utility. In *Proceedings of the 2006 Conference on ECAI 2006: 17th European Conference on Artificial Intelligence August 29 – September 1, 2006, Riva Del Garda, Italy*, pages 467–471, Amsterdam, The Netherlands, The Netherlands, 2006. IOS Press.

[15] Paul Weng. Axiomatic Foundations of Generalized Qualitative Utility. In *Multi-disciplinary International Workshop on Artificial Intelligence (MIWAI)*, Lecture Notes in Artificial Intelligence. Springer, dec 2013.